ED 393 407 IR 017 743

AUTHOR Tiemens, Robert K.
TITLE Analyzing the Content of Visual Messages: Methodological Considerations.
PUB DATE 93
NOTE 12p.; In: Verbo-Visual Literacy: Understanding and Applying New Educational Communication Media Technologies. Selected Readings from the Symposium of the International Visual Literacy Association (Delphi, Greece, June 25-29, 1993); see IR 017 742. Some figures may not reproduce clearly.
PUB TYPE Reports - Descriptive (141) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Coding; *Computer Software; *Content Analysis; Photography; *Production Techniques; Research Methodology; Research Problems; *Visual Literacy
IDENTIFIERS Information Value; Numeric Data; Variables; *Video Production; Visual Design; *Visual Imagery

ABSTRACT

The visual image is so information-rich that describing its content fully can be very problematic. Trying to codify or analyze the content in any systematic way often proves even more difficult. Nevertheless, reducing certain aspects of visual data to a set of numbers can be useful because it exposes characteristics of the message that might be obscured by other information. An interactive computer program called VIZ is being developed which allows researchers to input numerical values for different production techniques that are divided into seven variables: (1) shot length, or the duration for which a visual image appears on the screen; (2) style or pace of transitions between shots; (3) camera framing, from the close-up to the wide-angle shot; (4) the vertical camera angle; (5) vector orientation; (6) motion; and (7) shot content. It is hoped that VIZ's numerical profiles of screen images can render new insight into the nuances of visual information. Four figures illustrate the discussion. (Contains 16 references.) (BEW)

ANALYZING THE CONTENT OF VISUAL MESSAGES: METHODOLOGICAL CONSIDERATIONS

by
Robert K. Tiemens

The visual image is so rich, so abundantly loaded with information, that it is impossible to describe or codify its content in any complete way. The seemingly simple task of analyzing a 30-second television commercial which may contain hundreds of images (depending on how the unit of analysis is defined), can become a massive undertaking. Those who are engaged in the study of visual communication understand how impossible it is to construct a verbal description of a visual message. Visual images must be seen to be understood and appreciated.

The richness of the visual image can often conceal many of the relevant features contained in a visual message. Even the trained critic or observer sometimes overlooks the subtle use of production techniques because of the overwhelming amount of information that the visual message contains.

This paper is concerned with how the content of visual communication *text* can be systematically coded and analyzed. The paper sets out to achieve three objectives: (a) to identify specific production variables that shape the visual message--variables that give the message its form, (b) to operationally define and suggest a rationale for observing and analyzing those production variables, and (c) to describe the development of an interactive computer program that facilitates the systematic coding and analysis of visual data.

Systematic analysis of *visual* content has received limited attention (Avery & Tiemens, 1992; Barbatsis & Guy, 1991; Barker, 1985; Mayerle & Rarick, 1989; Porter, 1987, Tiemens, 1978, 1979; Tiemens, Sillars, Alexander, & Werling, 1988). Nevertheless, these studies demonstrate how empirical treatment of the data can enrich the analysis and interpretation of visual content.

One might suspect that to reduce a visual image to an array of empirical data, based on a discrete number of categories, would compound the difficult task of capturing the essence of the visual message. In fact, the opposite is true. By reducing the visual data to a set of numbers, we often expose characteristics of the message that are obscured by the massive amount of information contained in the images themselves.

There are two major obstacles in using a content-analysis approach to study visual images. First, the simple task of coding visual data can become extremely time-consuming due to the vast amount of information that visual images contain. The researcher must, therefore, avoid the temptation to code *everything* and be discerning in establishing the parameters that will guide the analysis. Second, even a reasonable number of coding categories can produce an immense amount of data, and the subsequent analysis of the data can be unwieldy.

VIZ: An Interactive Computerized Coding Program

VIZ is an interactive computer program designed to facilitate the coding of visual data. The program is still in the development stage, but several prototypes have been completed and tested.

VIZ presents a series of screens calling for the input of data for each variable. As the data are entered the program automatically advances to the next screen. That is, when a code "1" that represents a category is entered at the keyboard, the program automatically advances to the next screen. Thus, only one key stroke is needed to enter the datum which expedites greatly the process of coding.

As the data are entered, a description of the shot according to each code, is printed on the top third of the computer screen allowing the coder to confirm its accuracy. When all data have been entered and confirmed as being correct, the information for each shot is written to a sequential data file on the computer disk. This data file is formatted in such a way that it can be used in any standard statistical program.

Figure 1 shows a representative screen from the *VIZ* program. The example illustrates how the selection of *Medium Shot* as a descriptor for camera framing leads to a sub-category (presented through a pop-up window) that prompts the coder to enter the number of persons shown in the shot. The figure also illustrates how a summary of the coded information is recorded on the screen.

Table 1 outlines the basic categories that are included in the coding program. Each of these categories identify specific production variables that represent various formal qualities of a visual message. Each of these categories constitute a different *routine* in the *VIZ* coding program. A more complete discussion of each variable and the rationale for using each variable in the coding program follows.

Production Variables

Shot Length

The *shot* is the unit of analysis that is typically used to study visual content.[1] The relative duration (length) of a shot reveals two things: First, it indicates the degree to which visual details of an event are emphasized. The longer a visual image remains on the screen, the more dominant it becomes; and the cumulative time which an image is shown (combined for multiple shots) becomes a useful index for comparing and contrasting visual content.

Second, shot length determines the cutting rate of visual pacing of a mediate event. Zettl (1990) calls this manipulation of cutting rate tertiary motion and suggests that it can heighten the intensity of an event and influence the viewer's perception of whether the event is fast or slow. Penn (1971) has shown that cutting rate can affect the perceived potency of an activity of shot content, particularly when persons are depicted in the scene.

VIZ calculates the length of individual shots (in seconds) by simply subtracting the beginning time of one shot from the beginning time of the next shot. Thus, coding of shot length is automatically calculated from a minimum amount of information that the coder enters.[2] A useful technique to facilitate the coding of shot length is to insert or superimpose time-code markers on video tapes that are used for analysis. This

---

[1]A *shot* is usually defined in traditional film terms as the continuous, uninterrupted filming by the camera. However, this definition does not account for changes in the visual content that occur through movement of the objects within the frame or movement of the camera. Therefore, the VIZ program relies on an operational definition of shot as a change in the visual content of framing of the screen image.

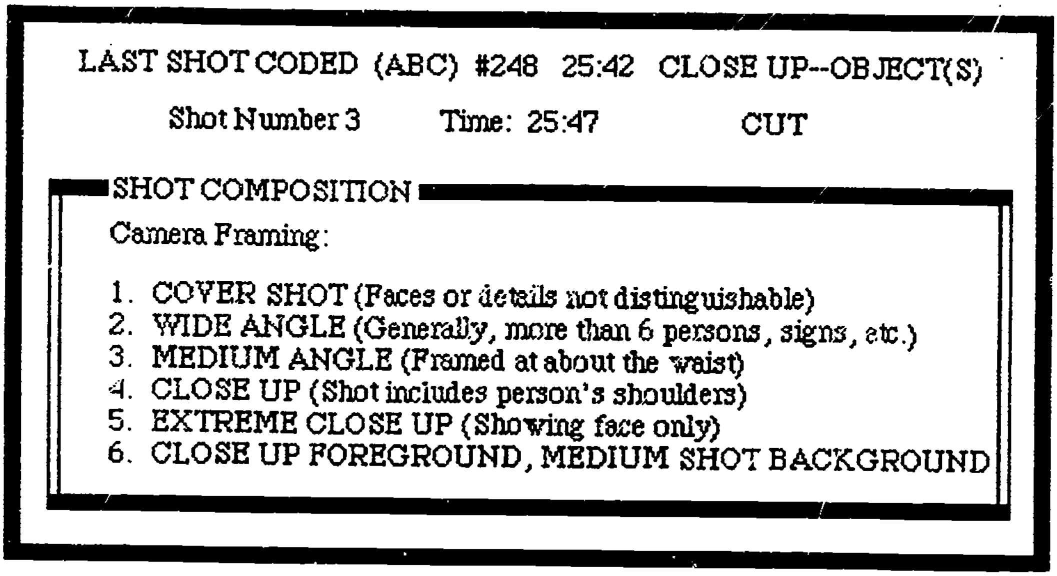
[2]Typically, shot length is recorded in number of seconds, though shorter units of measurement are possible and may be more suitable for coding some kinds of visual content. Some television spots incorporate shots that are less than 1/2 second in length. In analyzing the content of these messages it may be more useful to code shot length by the number of television frames. The present version of VIZ limits coding of shot length to seconds.

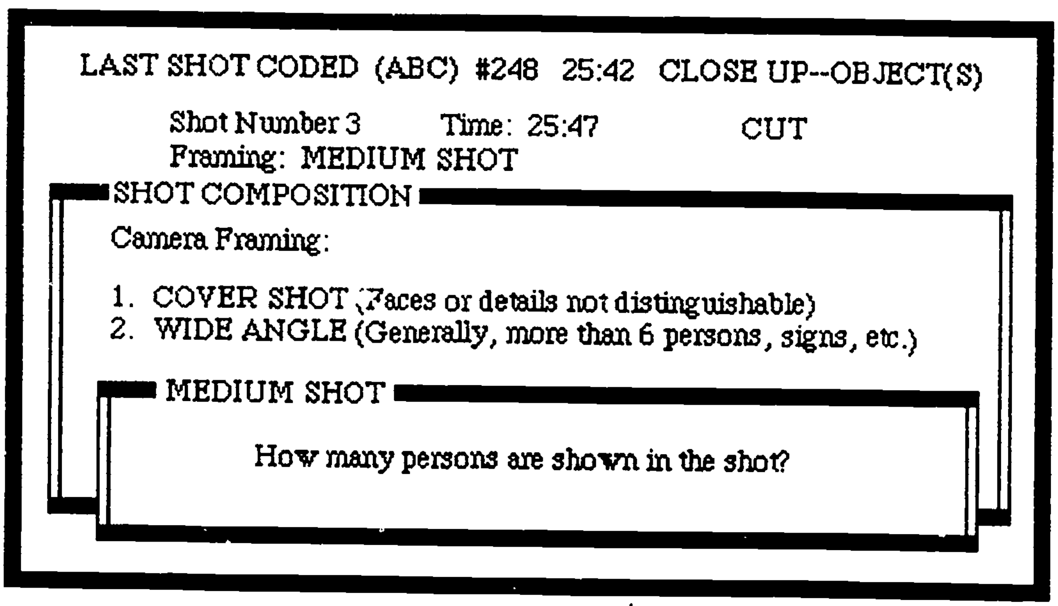
procedure allows for very precise coding of shot length (to 1/30th of a second, if desired.

Transition

Transitions define the relationship between shots; that is, how two visual images are related. Visual transitions, like shot length, influence the visual rhythm or pacing of the event. A *cut* produces a quick and abrupt change from one shot to the next. Combined with shots of short duration, cuts heighten the fast pacing or

Figure 1. Representative Screens from the *VIZ* Coding Program





This example illustrates how *VIZ* guides the researcher through the coding program. Coding of the visual image as a medium shot (code 3) prompts the coder to enter the number of persons shown in the shot. If the shot is coded as an extreme close up (code 5), the pop-up window would not appear since the ECU is defined as showing only a person's face. Instead the program automatically enters a "1" for the number of persons shown.

Table 1. Categories for Coding Visual Information

1. Identification of segment (e.g., shot number)
2. Shot transition
   a. Cut
   b. Dissolve
   c. Super Imposition or Long Dissolve
   d. Special effects (e.g., wipes, flips, etc.)
   e. Camera/lens movement, movement of subject or object
3. Length of segment (in seconds)
4. Camera framing
   a. Cover shot
   b. Wide angle shot
   c. Medium shot
      (1) Number of persons shown in the shot
   d. Close up
      (1) Number of persons shown in the shot
   e. Extreme close up (only one person shown)
   f. Close up foreground, Medium shot background
5. Camera treatment
   a. Objective camera treatment
   b. Subjective camera treatment
6. Vertical camera angle of person(s) shown
7. Vector orientation (camera position relative to person(s) shown)
8. Camera/subject motion
   a. Static (no motion)
   b. Subject movement
      (1) Movement along the Z-axis
      (2) Movement only along the X-Y axis
      (3) Direction of movement not relevant to the shot
   c. Camera movement
      (1) Zoom in/out
      (2) Dolly in/out
         (a) Uses Z-axis staging
         (b) Does NOT use Z-axis staging
         (c) Z-axis not relevant to the shot
      (3) Pan/Tilt
      (4) Boom, truck, etc.
   d. Subject + Camera movement
      (1) Uses Z-axis staging
      (2) Does NOT use Z-axis staging
      (3) Z-axis not relevant to the shot
8. Shot content
   a. Primary Category
      (1) Sub-category 1
         (a) Sub-category 2
            (*1*) Sub-category 3
   b. Etc.
   c. Etc.

tertiary motion of an event. *Dissolves*, on the other hand, have a more fluid quality and tend to slow the pace and convey a more relaxed mood. In addition, dissolves create a momentary superimposition which visually blends two images together.

Long dissolves (i.e., 2 seconds or longer) strengthen the connection or visual association between two shots. As one shot slowly blends into the next, the visual images are fused together into a composite image, emphasizing the relationship between the visual content of the two shots. *VIZ* records super impositions and long dissolves as a separate category so that the researcher can examine these instances more closely.

Wipes, flips, rotations, squeeze zooms, and other electronic transitions can impart special connotations of how two images are related. An electronic flip, for example, could show a contrast between visual images suggesting that they represent dichotomies or opposing points of view. Similarly, Zettl (1990) demonstrates how a wipe presents the image of the new shot pushing the old one off the screen. In this manner, the wipe can suggest the replacement of one concept by another, visualized by *pushing away* one image and *replacing* it with another.

Camera movement, lens movement, and *subject movement* often alter the visual content or makeup of the screen image. For example, a closeup image of a person or object may be interpreted differently when the lens zooms out to reveal the context of that image or to show the juxtaposition of another person or object. *VIZ* treats camera movement or lens movement as a traditional device, marking the end of one shot and the beginning of another. Of course, these changes can be ignored by the researcher when analyzing the raw data produced by the coding program.

Entering the choice of camera or subject movement also cause the program to branch to a sub-routine that allows for the coding of more detailed information (see description under the subheading of Motion).

Camera Framing

Camera framing alters the relative size of an image within the defined space of the television frame, thereby influencing the visual emphasis or intensity of the image. *Closeups* create greater intensity and give the image greater importance by directing the viewer's attention to eventful visual detail. *Wide angle* shots, on the other hand, diminish the relative size of persons or objects shown on the screen and thus de-emphasize their visual importance.

Labels used to describe camera framing (e.g., close-up, medium shot, wide-angle shot, etc.) are relative terms and are subject to individual interpretation. Therefore, in coding such data it is useful to have more complete descriptions so that camera framing is consistently defined for each shot. *VIZ* includes the following descriptors to aid coders' judgments:

Cover shot	(faces or details not distinguishable)
Wide angle	(generally more than 6 persons shown on the screen)
Medium shot	(person shown on camera framed at about the waist)
Closeup	(shot of person shows his or her shoulders)
Extreme Closeup	(shows only the person's face or less)
Closeup foreground, Medium shot background	

If the shot is coded as a medium shot or closer, *VIZ* prompts the coder to enter the number of persons shown on camera. This becomes useful for identifying two-shots or other groupings that might be relevant to the visual analysis.

Alternatively, visual aids such as the one shown in Figure 2, can be used to illustrate more explicitly variations in camera framing. Aids such as these reduce the ambiguity that is inherent in the traditional nomenclature of camera composition, and help to increase the reliability of coders' judgements.

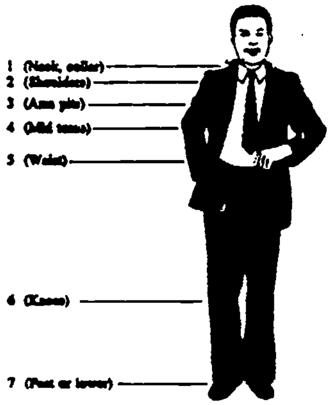


Figure 2. Visual Diagram to Facilitate Coding of Camera Framing

Camera Treatment

Most often the function of the camera is to observe the event, in which case it is used as an *objective* camera. The viewpoint of an objective camera is as an unseen observer and the persons seen on the screen seem unaware of its presence. *Subjective* camera treatment occurs when the person or persons shown on screen look at or talk directly to the camera. In this case, the camera no longer observes the event but rather participates in the event. The camera, and consequently the viewer, becomes more directly involved in the screen event. Objective/subjective camera treatment is an optional coding category in the *VIZ* program.

Vertical Camera Angle

Armer (1986) describes how the vertical camera angle influences our perception of a person shown on the screen:

> When we look down on someone, figuratively as well as literally, we place him or her in an inferior position. He or she becomes subordinate, recessive, smaller than we. . . .We look up to people we respect, who occupy a higher position in society, who tower above us intellectually or professionally. Similarly, when a director places the camera below eye level, looking up at a character, that character assumes a position of dominance, of strength, of importance. (pp. 183-184)

There is considerable empirical support for Armer's explanation (Kepplinger & Donsbach, 1990; Mandell & Shaw, 1973; McCain, Chilberg, & Wakshlag, 1977; Tiemens, 1970); however, as Kepplinger (1991) points out:

> . . .research in this area does not show a consistent effect for specific camera angles. . . .The codes activated by camera angle are no doubt interacting with other codes that are simultaneously processed. *Together* they construct the meaning and effect of a particular shot. (p. 32)

Vertical camera angle becomes particularly relevant when two persons are shown on camera, as in an over-the-shoulder shot. In such instances the vertical positioning of the camera will determine the relative screen placement of the two persons shown in the frame and thus give one person a visual advantage over the other. The person whose image is higher in the frame generally takes on a greater dominance and creates the illusion that he or she is looking down (literally and figuratively) on the other person. This phenomenon was clearly evident in the 1976 and 1988 presidential debates (Tiemens, 1978, 1989).

The *VIZ* coding program offers the following choices for coding vertical camera angle of a shot:

High camera angle	(Camera looking DOWN at subject)
Medium camera angle	(Camera positioned at EYELEVEL)
Low camera angle	(Camera looking UP at subject)
Camera angle not applicable	

Vector Orientation

The degree to which televised events are visually unified and coherent is the result of how vector combinations maintain an appropriate pattern of screen direction. For example, a closeup of a person facing screen-left juxtaposed with a shot of another person facing the screen-right show *converging* vectors and suggest that the two are looking at or interacting with one another. If both shots show the two individuals facing the same direction, it suggests that one person is looking away from the other (*continuing vectors*). Zettl (1990) illustrates how, in a public speaking event, the vector combination maintains an expected speaker-audience relationship if shots of the speaker and the audience produce a converging vector combination.

The *VIZ* program treats *vector orientation* as the relationship between camera position and the person(s) shown on camera. The program presents a graphic illustration to facilitate the coding of this data (see Figure 3). By analyzing the vector orientations of juxtaposed shots, the researcher can establish the instances in which a televised presentation uses converging or continuing vectors.

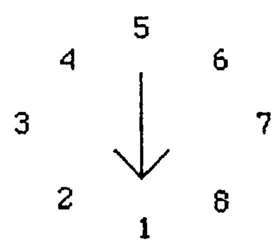
Figure 3

LAST SHOT CODED (ABC) #121 23:14 CLOSE UP-- DELEGATES/ FLOOR AUDIENCE

Shot Number 13 Time: 23:21 DISSOLVE
Framing: MEDIUM SHOT 2 PERSONS SHOWN
Shot Content: DELEGATES/FLOOR AUDIENCE
Demographics: WHITE, MALE, 35 YEARS AND YOUNGER
HIGH CAMERA ANGLE (Camera looking DOWN at subject)

CAMERA POSITION

1. Front
2. Left Front
3. Left Side
4. Left Back
5. Back
6. Right Back
7. Right Side
8. Right Front



Graphic aids can be presented on the screen to facilitate the coding of data. In this example, the diagram illustrates the camera position (relative to the subject being show) to code vector orientation.

Motion

Motion is a particularly relevant variable because it attracts our attention more than any other visual element. Thus, we usually attend more to what is happening on the screen when the image is moving. Of course, this could be detrimental if the motion detracts from the point of the message; i.e., it can create communication noise.

Zettl (1990) classifies motion according to three distinct types: (1) *primary motion*, the movement of subjects or objects in front of the camera, (2) *secondary motion*, movement induced by the camera or lens, and (3) *tertiary motion*, the movement and rhythm established through editing.[3]

An important element of primary and secondary motion is the axis along which the motion occurs. Zettl (1990) notes that motion along the z-axis (toward or away from the camera or along the axis that the camera is pointing) is particularly advantageous on the small television screen and creates one of the strongest indicators of depth. Thus, any coding scheme that includes motion as a primary variable should classify that motion according to whether it or not it occurs along the z-axis.

Figure 4 shows how the coding program accommodates the classification of different levels of camera and subject movement. If subject movement or camera movement are important variables the program cues the coder for more detailed descriptions of the movement, including whether or not the movement occurs along the z-axis.

Shot Content

Shot content is a crucial variable in visual analyses. The relevancc of other variables used in the coding scheme is often found in how they interact with, or are used to define the visual content of a shot. The categories that one can use to classify visual content are endless and must be chosen on the basis of the specific research questions.

The coding program allows for eight general categories of visual content. Beyond these categories, three additional sub levels of categories can be included with a maximum of six categories in each sub level. Thus, if *human subject* represents a general choice for shot content, sub levels may include *gender*, *race*, or *age* of the person shown on camera.

Practical Considerations and Advantages

An interactive computer program, such as the one described here, offers a number of advantages for content analyses of visual media. An enumeration of some advantages follows.

A primary advantage of the technique is that the process of coding is much faster. Because each category is represented by a single screen (or pop-up windows within the screen), there is no need for complicated coding sheets. Entering a code for each variable is accomplished by a single key stroke and the laborious task of writing numbers by hand on a coding sheet is eliminated. When a code is entered, the program automatically advances to the next category.

The ease and efficiency of entering data are also facilitated by having the program enter some data automatically. For example, each unit of analysis (in this case a shot number) is identified and entered to the data file automatically. Similarly, the variable of shot length is calculated by the program and entered into the data file automatically; the coder

---

[3]Tertiary motion is operationalized basically by the length of shots in a video segment. For that reason I have included the concept of tertiary motion under the general heading of shot length.

Figure 4

LAST SHOT CODED (ABC) #248 25:42 CLOSE UP--OBJECT(S)

Coding Shot #249 at 25:47
Medium Shot 2 persons Female
High Camera Angle Front left

MOTION WITHIN THE FRAME
1. Static (no motion)
2. Subject Movement
3. Camera Movement
4. Subject Movement + Camera Movement

LAST SHOT CODED (ABC) #248 25:42 CLOSE UP--OBJECT(S)
Coding Shot #249 at 25:47
Medium Shot 2 persons Female
High Camera Angle Front left

MOTION WITHIN THE FRAME
1. Static (no motion)
2. Subject Movement
3. Camera Movement

SUBJECT + CAMERA MOVEMENT
1. Zoom/dolly in
2. Zoom/dolly out
3. Pan
4. Tilt
5. Other (e.g., boom up/down, truck, etc.)

LAST SHOT CODED (ABC) #248 25:42 CLOSE UP--OBJECT(S)
Coding Shot #249 at 25:47
Medium Shot 2 persons Female
High Camera Angle Front left

MOTION WITHIN THE FRAME
1. Static (no motion)
2. Subject Movement
3. Camera Movement

SUBJECT + CAMERA MOVEMENT
1. Zoom/dolly in

ZOOM/DOLLY OUT
1. Utilizes z-axis staging
2. Does NOT utilize z-zxis staging
3. Z-zxis not relevant to shot

Sub-categories of coded information are presented by windows. In this example, three levels of information about *motion* are presented. Organizing the categories in this fashion helps the coder to proceed in a logical manner from one step to the next.

only needs to enter the ending time for each shot.[4]

VIZ can be programmed to skip categories which are not relevant and automatically enter a zero data code. For example, if a shot is coded as a *cover shot* the program will skip those categories that call for close-up detail (e.g., number of persons in the shot; gender, race, or age of persons shown in the slot, vertical camera angle, etc.).

The program prohibits codes which are *out of range* from being entered, thus reducing the likelihood of miscoding the visual material. For example, if the maximum data code is *6* and the coder enters a higher number, the program will sound a beep, alerting the coder that a different code must be entered. Similarly, the time marking the ending of each shot must be greater than the ending of the previous shot, thus reducing the chances of entering wrong data.

Internal checks for reliability or consistency of the data are built into the program whenever possible. For example, if a medium shot is coded as showing more than six persons within the frame, the program will remind the coder that the shot should be classified as a wide angle shot and offer the opportunity to recode the shot.

Finally, a major advantage is that the data are written directly to a disk file which can be translated to a spreadsheet program or as a data file for a statistical program such as SAS, SPSS, MINITAB, etc. This avoids the step of writing the data by hand and then rekeying it for computer analysis. None of the data is coded by hand or in written form which further diminishes the likelihood of coding errors.

[4]The beginning of time of each slot is automatically recorded as the time at which the previous shot ended, but the coder does have the option of entering a different beginning time.

Finally, we must recognize that an empirical analysis of visual content is simply one approach toward interpreting the *text* of televised messages and cannot stand by itself. Empirical analysis of visual content must be accompanied by other, more qualitative, methods of interpretation. It has the advantage, however, of rendering a complete and objective *picture* of what the visual images contain. A tremendous strength of the content analysis approach, as suggested by this paper, can render tremendous insights by *revealing* nuances that are otherwise overlooked. As such, it can enhance our understanding and interpretation of visual messages.

References

Armer, A. (1986). *Directing television and film*. Belmont, CA: Wadsworth.

Avery, R. K., & Tiemens, R. K. (1992). Tracking the eye: An investigation of a combat photographer's visual odyssey. *Feedback, 33*(1), 12-15.

Barbatsis, G., & Guy, Y. (1991). Analyzing meaning in form: Soap opera's compositional construction of "realness." *Journal of Broadcasting and Electronic Media, 35*, 59-74.

Barker, D. (1987). Television production techniques as communication. *Critical Studies in Mass Communication, 2*, 234-246.

Kepplinger, H. M. (1991). The impact of presentation techniques: Theoretical aspects and empirical findings. In F. Biocca (Ed.), *Television and political advertising. Vol. 1* (pp. 173-194). Hillsdale, NJ: Lawrence Erlbaum Associates.

Kepplinger, H. M., & Donsbach, W. (1990). The impact of camera perspectives on the perception of a speaker. *Studies in Educational Evaluation, 16*, 133-156.

Mandell, L. M., & Shaw, D. L. (1973). Judging people in the news--unconsciously: Effect of camera angle and bodily activity. *Journal of Broadcasting, 17*, 353-362.

Mayerle, J., & Rarick, D. (1989). The image of education in primetime network television series 1948-1988. *Journal of Broadcasting and Electronic Media, 33*, 139-157.

McCain, T. A., Chilberg, J., & Wakshlag, J. (1977). The effect of camera angle on source credibility and attraction. *Journal of Broadcasting, 21*, 35-46.

Penn, R. (1971). Effects of motion and cutting-rate in motion pictures. *AV Communication Review, 19*, 29-51.

Porter, M. J. (1987). A comparative analysis of directing styles in Hill Street Blues. *Journal of Broadcasting and Electronic Media, 31*, 323-334.

Tiemens, R. K. (1970). Some relationships of camera angle to communicator credibility. *Journal of Broadcasting, 14*, 483-490.

Tiemens, R. K. (1978). Presidential debates: An analysis of visual content. *Communication Monographs, 45*, 362-370.

Tiemens, R. K. (1979). The visual context of argument: An analysis of the September 25, 1988 presidential debate. In B. Grombeck (Ed.), *Spheres of argument* (pp. 140-146). Annandale, VA: Speech Communication Association.

Tiemens, R. K., Sillars, M. O., Alexander, D. C., & Werling, D. S. (1988). Television coverage of Jesse Jackson's speech to the 1984 Democratic National Convention. *Journal of Broadcasting and Electronic Media, 32*, 1-22.

Zettl, H. (1990). *Sight, sound, motion: Applied media aesthetics* (2nd ed.). Belmont, CA: Wadsworth.